On the Efficient Allocation of Resources for Hypothesis Evaluation: A Statistical Approach

Steve Chien, Jonathan Gratch, and Michael Burl

Manuscript received March 8, 1994, revised Feb. 2, 1995.

S. Chien and M. Burl are affiliated with the Jet Propulsion Laboratory, MS 525-3660, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA, 91109-8099. Email: {chien, burl}@aig.jpl.nasa.gov.

J. Gratch is affiliated with the Beckman Institute, University of Illinois, 405 N. Mathews Ave., Urbana, IL, 61801. Email: gratch@cs.uiuc.edu

IEEECS Log Number P95085

Abstract

This paper considers the decision-making problem of selecting a strategy from a set of alternatives on the basis of incomplete information (e.g., a finite number of observations). At any time the system can adopt a particular strategy or decide to gather additional information at some cost. Balancing the expected utility of the new information against the cost of acquiring the information is the central problem we address.

In our approach, the cost and utility of applying a particular strategy to a given problem are represented as random variables from a parametric distribution. By observing the performance of each strategy on a randomly selected sample of problems, we can use parameter estimation techniques to infer statistical models of performance on the general population of problems. These models can then be used to estimate: (1) the utility and cost of acquiring additional information; and (2) the desirability of selecting a particular strategy from a set of choices. Empirical results are presented that demonstrate the effectiveness of the hypothesis evaluation techniques for tuning system parameters in a NASA antenna scheduling application.

Keywords

machine learning, the utility problem, planning and scheduling, parameter estimation, adaptive problem-solving

I. Introduction

In machine learning and basic decision-making in AI, a system must reason about alternative courses of action in the absence of perfect information; frequently, the expected utility of the information to be acquired must be balanced against the cost of acquiring the information. When one wishes some sort of statistical guarantees on the (local) optimality of the choice and/or the guarantee of rationality, a statistical decision theoretic framework is useful. This problem of decision-making with incomplete information and information costs can be analyzed in two parts:

- How much information is enough? At what point do we have adequate information to select one of the alternatives?
- If one wishes to acquire more information, which information will allow us to make the best possible decision at hand while minimizing information costs?

Possible solutions to this decision-making quandary depend on the context in which the decision is being made. This paper focuses on an abstract class of decision problems called *hypothesis selection problems* that arise in many contexts in machine learning. These problems arise when one must select the best hypothesis (such as hypothesized concept description, or a hypothesized problem-solving heuristic) from a set, given its performance over some training data. For example, in adaptive problem solving a learning algorithm must select, from a set of possible control strategies, one that most enhances problem solving performance [1], [2]. In inductive learning there are two issues which are naturally seen as hypothesis selection problems: the attribute selection problem consists of selecting one of a set of attributes based on information gain [3], [4]; and the model selection problem consists of selecting one of a set of learned models (e.g. pruned decision trees) based on their classification accuracy [5]. Although hypothesis selection problems occur in many contexts, in this article we will use the terminology appropriate for adaptive problem-solving - so that acquiring additional information corresponds to solving problems with a particular problem solving strategy and selecting a problem-solving strategy with high expected utility is the goal.

Solving hypothesis selection problems may involve significant investment of resources. There may be monetary cost in obtaining training data and computational cost in processing it. Usually this cost is addressed by informal or intuitive judgements rather than a rational analysis of the costs and benefits involved. This paper introduces two general methods for solving hypothesis selection algorithms efficiently and each method can be augmented with rational analysis to minimize the total cost of selecting a hypothesis. The first method, called *interval-based selection*, involves quantifying the uncertainty in competing hypotheses by using the statistical confidence that one hypothesis is better than another hypothesis. In this approach the system allocates examples to show that one hypothesis dominates all the other hypotheses with the specified confidence. These methods also rely upon an indifference parameter – if two hypotheses differ in performance by less than this amount, either is acceptable. [1]

The second method uses the decision theoretic concept of *expected loss* [7], [8], which measures the probability of making a less preferable decision weighted by the lost utility with respect to the alternative choice. In the expected loss approach, the system acquires information until the expected loss is reduced below some specified threshold. This

[1] This formalism is analogous to the PAC [6] framework – "probably" "approximately" "correct" maps onto "probably" "close to" "highest expected utility".

approach has the added benefit of not attempting to distinguish among two hypotheses with similar means and low variances (e.g., it recognizes indifference without a separate indifference parameter).

For both the interval-based and expected loss approaches, when comparing among more than two alternatives, one is comparing the utility of the "best" hypothesis to the other possible hypotheses. Since there are multiple comparisons, the estimate for the overall error in the final conclusion (selection of a best hypothesis) is based upon the errors associated with multiple smaller conclusions. In both the interval-based and expected loss approaches, it is possible to improve performance by rationally allocating varying amounts of error to each of the smaller conclusions. Hence, there are four algorithms we consider: interval-based with equal error allocation, interval-based with unequal (rational) error allocation, expected loss with equal error allocation, and expected loss with unequal (rational) allocation.

The rest of this paper is organized as follows. Section 2 describes the general hypothesis evaluation problem and frames the problem as statistical parameter estimation. Section 3 describes the confidence interval approach. Section 4 describes the expected loss approach and Section 5 describes an empirical evaluation of these techniques using synthetic and real-world scheduling data. Section 6 summarizes the principal points of this paper.

II. The Hypothesis Evaluation Problem

Hypothesis evaluation is the problem of selecting one of a set of hypotheses which, with high probability, is close to the best. We adopt a parametric statistical approach to this problem. Typically we have a set of problems D (planning problems, exemplars to classify, etc.). Any particular problem d is selected from this set with probability $P_D(d)$. We also have available a set of k potential alternative strategies $H_1, \ldots, H_k$, for solving problems. Each hypothesized strategy H_i has associated with it an unknown utility distribution U_i describing its quality, and an unknown cost distribution C_i describing the cost to process examples. Both of these are are induced by the probability distribution over D.[2] The

[2] Considerable work has been devoted to speedup learning, in which U_i and C_i often are inversely related. For example, in speedup learning one might use U = -C. In other work the utility of a solution might relate to the quality of the overall plan or schedule produced [9], [10], [11].

desired outcome of the hypothesis evaluation problem is to select a hypothesis H_{best} which has the highest (or close to highest) expected utility [3].

Although the distributions U_i and C_i are unknown, the decision-making system can infer information about these distributions by observing the behavior of strategy H_i on problems drawn from D. Thus, the system can choose between *acquiring more information* – acquiring another sample from U_i with cost drawn from C_i or *adopting a hypothesis strategy* H_i (the same question as in the introduction).

Our general approach to this problem consists of two parts: parameter estimation and hypothesis evaluation. In parameter estimation the underlying distributions of expected utility and expected cost are assumed to be of a particular form (e.g., normal, student T, etc.) reducing the problem to one of estimating parameters such as mean and variance from behavior on sample problems. In hypothesis evaluation, decision rules to decide how much information is enough, and how to acquire information are formulated based upon estimated parameters. As the result of applying these decision rules, the system may decide to gather additional information (samples), in which case it faces the decision between acquiring information and stopping again. This process continues until the system determines it has acquired enough information.

For purposes of estimating the expected value of these distributions we assume that U_i and C_i are jointly normally distributed (sometimes called gaussian) random variables with unknown means and unknown *general* covariance. The assumption of normality is quite reasonable as the estimated expected value of an arbitrary distribution is approximately normally distributed (a consequence of the Central Limit Theorem [13]). Confidence intervals regarding the true mean can be computed from the sample mean, sample variance, and number of samples. More concretely, one can show that the difference between the observed sample mean and true mean is normally distributed with 0 mean and $\frac{1}{n}$ times the variance of the initial distribution, e.g. $\hat{\mu} - \mu \sim N(0, \frac{\sigma^2}{n})$ [14].

Given the assumption of normality we can also conclude that the differential distribution (the distribution of the difference in utility between any two strategies) is normally distributed. This property is important in that it allows us to determine that one strat-

[3] Alternative castings of the problem might also impose requirements on the variance of the selected distribution (e.g., [12]).

egy is better than (or roughly equivalent to) another strategy in expected utility by only maintaining information about the differential distributions. This simplifies some of the mathematics. For example, in many applications the performance of different strategies may be highly correlated (e.g., when strategies are small modifications of some common ancestor). Using the differential distributions encodes this correlational information without the need for explicitly computing covariance estimates.[4]

A. Other Approaches

Our approach to hypothesis evaluation is related to several other methods in the machine learning and statistics literature. Standard machine learning approaches do not provide bounds on the quality of the selected hypothesis, and thus do not fit into our conception of hypothesis evaluation. However, hypothesis evaluation proper has been studied extensively in computational learning theory. The thrust of that community has focused on the question of whether hypothesis selection is possible in the worst possible circumstances (and thus avoids parametric approaches); however, we are concerned with algorithms that are highly efficient in practice. The closest approaches to ours from the computational learning theory community are [2], [17].

In the statistics literature, hypothesis evaluation problems are refered to as ranking and selection problems [18]. In their terminology we are studying sequential elimination selection procedures [19]. Our work differs from this literature in that our approach is more general. Standard selection techniques make restrictive assumptions about the variances of the utility distributions. We allow the utility distributions to be correlated and have unequal (finite) variances. However, we give up the strong correctness proofs provided by these statistical techniques. Our techniques are heuristic and we provide only mathematically plausible and empirical arguments for their correctness. The approach closest to ours in generality is a machine learning technique proposed by Moore and Lee [5].

Our approach extends both learning theory and statistical approaches in that we account for the cost of obtaining data. Typically hypothesis selection approaches only attempt to minimize the overall number of examples. We extend these approaches to account for

[4] This technique is known as *blocking* in the statistical literature, see p. 299-300 of [15], and the method of *common random variables* in the simulation literature [16].

situations where the cost of evaluating different hypotheses substantially differs.

B. Notation

Throughout this paper we use the following notation:

- U_i is the utility distribution for the hypothesis strategy H_i
- C_i is the cost distribution for the hypothesis strategy H_i
- μ_i is the true mean for the variable U_i
- $\overline{U}_i$ is the sample mean for the variable U_i
- σ_i is the true standard deviation for the variable U_i
- S_i is the sample standard deviation for U_i
- $\overline{C}_i$ is the sample mean for the variable C_i
- U_{i-j} is the variable for the distribution computed by taking the utility of H_i minus the utility of H_j both solving the same problem. Note that this distribution is Gaussian (normal) if U_i and U_j are jointly gaussian even if U_i and U_j are not independent. μ_{i-j}, $\overline{U}_{i-j}$, σ_{i-j}, and S_{i-j} are analogously defined.

We also define functions to allow computation of probabilities of normally distributed variables. The probability that a random variable y has a value in the interval (a,b) given that the variable is normally distributed with mean μ and standard deviation σ is

$$\Phi(a,b;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma}\int_a^b e^{-0.5(\frac{y-\mu}{\sigma})^2}dy$$

For the standard normal distribution with mean μ=0 and standard deviation σ=1, we use the specialized notation:

$$\Phi(a,b) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-0.5y^2}dy$$

III. The Interval-based Approach

The confidence interval-based approach depends on a confidence parameter γ and an indifference parameter ϵ. This approach attempts to show that with confidence γ there is a hypothesis strategy H_i such that for every other hypothesis strategy H_j, either: a) $E[U_{i-j}] > 0$ or b) $|E[U_{i-j}]| < \epsilon$. Intuitively, if such an H_i can be found it should be adopted because for every other hypothesis strategy H_j, with confidence γ, either H_i is better than H_j (dominance) or H_i and H_j are close enough so that we do not care (indifference). This intuitive description will be further elaborated in the following paragraphs.

Consider two of the hypothesis strategies being evaluated H_i and H_j. Under the assumption that U_i and U_j are jointly normally distributed, the difference U_{i-j} is normally distributed. Hence, analyzing the difference U_{i-j} and computing the confidence that $\mu_{i-j} > 0$

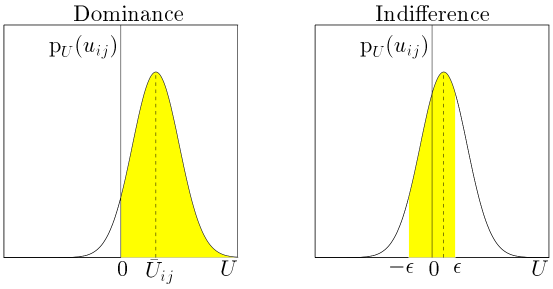

Fig. 1. Dominance and Indifference Calculations

gives the confidence that H_i dominates H_j. To represent the confidence in this pairwise comparison of U_i and U_j we use the variable γ^*.

To compute the confidence that $\mu_{i-j} > 0$ we adapt a method for computing confidence intervals for the mean of a normal distribution with unknown variance from [14]. However, our application differs from the standard confidence interval calculation as follows. In the standard problem, one is given a confidence level γ^*, and the task is to compute an interval such that the true mean lies in the interval with confidence γ^*. In our case, we are given the interval, and we wish to compute the confidence that the mean lies within the interval. Thus, since $\mu - \overline{U} \sim N(0, \frac{\sigma^2}{n})$ for this difference distribution, and the confidence that $\mu - \overline{U}$ is in some interval is simply the integral of the normal curve for that interval, these assumptions result in the following formula (shown graphically in Figure 1):

$$\gamma^* = \Phi(0, \infty; \overline{U}_{i-j}, \frac{S^2_{i-j}}{n}) = \Phi(-\overline{U}_{i-j}\frac{\sqrt{n}}{S_{i-j}}, \infty)$$

To handle the case of indifference pruning, the confidence that $-\epsilon < \mu_{i-j} < \epsilon$ can be computed similarly to the method described above yielding the following formula (shown graphically in Figure 1):

$$\gamma^* = \Phi(-\epsilon, \epsilon; \overline{U_{i-j}}, \frac{S^2_{i-j}}{n}) = \Phi(\frac{(\overline{U}_{i-j} - \epsilon)\sqrt{n}}{S_{i-j}}, \frac{(\overline{U}_{i-j} + \epsilon)\sqrt{n}}{S_{i-j}})$$

This can be interpreted using the confidence interval stopping criterion as follows. In the first case γ^* indicates our confidence in the hypothesis that the mean of the distribution U_i is greater than the mean of the distribution U_j, thus we prefer H_i over H_j (dominance). In the second case the difference between the means of U_i and U_j is less than ϵ with confidence γ^*, thus H_i and H_j are not worth distinguishing (indifference). If $\overline{U}_i < 0$, then H_j appears to be superior to H_i so we should be focusing on H_j and not H_i.

One complication is that in a general hypothesis evaluation problem, one is selecting from $k > 2$ hypotheses. Thus, for the interval-based approaches, one is comparing one hypothesis H_{high} (believed to be the best) against the other $k-1$ hypotheses. Thus the confidence γ of the overall decision depends upon the confidences γ^* in the individual $k-1$ comparisons. If we presume a pessimistic accumulation of error, we might project that the errors would add — requiring that the sum of the $k-1$ errors add to less than γ.

Evenly distributing the error indicates that the individual confidences must be: $\gamma^* = 1 - \frac{1-\gamma}{k-1}$ (confidence equation 1). Unfortunately, in the worst case, for k strategies, the choice of the final selection may depend upon more than $k-1$ pairwise comparisons. Consider the case where the focus strategy H_{high} changes frequently while attempting to find a best strategy. Indeed, in the worst case, the final selection would depend upon all of the pairwise combinations of selections of two of the k strategies (due to shifting of the focus hypothesis strategy). This is simply k choose 2 or $k(k-1)/2$. Thus, in the worst case, for the equal distribution of errors premise, the individual confidences must be: $\gamma^* = 1 - \frac{2(1-\gamma)}{k(k-1)}$ (confidence equation 2).

However, typically one samples evenly from all of the distributions n_0 samples before one chooses a focus strategy. If n_0 is large enough such that the focus strategy H_{High} changes rarely, the overall confidence will more closely resemble the linear relationship described in confidence equation 1. Indeed, if the errors tend to cancel each other, even this linear summation of errors will be an overestimate of the actual error[5].

However, equal error allocation does not take advantage of the fact that reducing the error in some of the terms may be easier than in others. Pertaining to this issue we first outline an algorithm called STOP1 which distributes the error evenly, then show a variation on this basic algorithm STOP2 which accounts for the varying difficulty in reducing the error in each of the terms and takes into account the varying cost of sampling from each of the distributions.

A. The STOP1 Algorithm

The STOP1 algorithm can be described as follows. Let T be the set of hypothesis strategies $H_1, \ldots, H_k$. Sample from each of the utility distributions $U_1, \ldots, U_k$ some default number of samples n_0. Let H_i be the strategy in T which has the highest sample mean for U_i so far (hereafter called the focus strategy H_{high}) . For each strategy H_j in T, if

[5] For a further discussion of this issue see [20] p. 18-19

U_j is in the interval $-\epsilon, \epsilon$, attempt to show indifference. If not, attempt to show that H_i dominates H_j.

Indifference is shown as follows. Compute the confidence that the true mean μ_{i-j} of U_{j-i} lies within the interval $-\epsilon, \epsilon$. If this confidence is greater than γ^* then indifference has been shown. Else, sample from U_i and U_j as necessary until either: (1) the confidence that μ_{j-i} is within the interval $-\epsilon, \epsilon$ is greater than γ^* or (2) $\overline{U}_{j-i}$ goes above ϵ or below $-\epsilon$. If $\overline{U}_{j-i}$ goes above ϵ, U_j now has a higher sample mean than U_i by a significant amount so that we should make H_j the target hypothesis and proceed. If $\overline{U}_{j-i} < -\epsilon$, H_j looks significantly worse than H_i so that we should attempt to show that H_i dominates H_j.

Dominance is shown similarly. Compute the confidence that $\mu_{i-j} > 0$. If this confidence is greater than γ^* we have shown dominance; otherwise sample from U_i and U_j as necessary until either the confidence becomes greater than γ^* or U_{j-i} goes below ϵ. In this case, we might attempt to show indifference among H_i and H_j.

It is worth noting that sometimes when $\overline{U}_{j-i}$ is in the interval $-\epsilon, \epsilon$, there is more confidence in the claim that H_i dominates H_j than in the claim that H_i and H_j are indifferent. It is unclear whether a closed form exists that can be used to determine whether dominance or indifference has higher confidence. We avoid this problem by computing both the dominance and indifference and using the higher of the two confidences.

```
STOP1 ALGORITHM:
let T = $H_1, \ldots, H_k$
let $\gamma^* = 1 - (1 - \gamma)/(k - 1)$
solve $n_0$ problems with each strategy in T and compute U statistics
let $H_{High}$ be the strategy in T with the highest $\overline{U}$
LOOP1
    let $H_{High}$ be the strategy in T with the highest
        sample mean for $U_{high}$
    if for every $H_j$ in T one of the following holds
        $H_{high}$ dominates $H_j$ with confidence $\gamma^*$
        $H_{high}$ and $H_j$ are ambivalent with confidence $\gamma^*$
    then return $H_{High}$
    else select a strategy $H_j$ such that neither
        $U_{high}$ dominates $U_j$ with confidence $\gamma^*$
        nor $U_{high}$ and $U_j$ are ambivalent with confidence $\gamma^*$
    generate data for the distribution $U_{high-j}$
    recompute U statistics
CONTINUE WITH LOOP1
```

Note that the algorithm has been simplified for purposes of clarity. A realistic implementation would temporarily classify the strategies into indifference and dominance classes when confidence has been shown. When H_{high} changes, these strategies must be returned to the unknown pool because they must be compared to the new H_{high}.

B. The STOP2 Algorithm

The STOP2 algorithm differs from the STOP1 algorithm in that it accounts for two factors ignored in the STOP1 approach. First, depending upon the sample variances and sample means of the individual U_{j-i} distributions, examples allocated to the distributions will have different effects on improving the confidence in a pairwise dominance or indifference relation. Second, the cost of acquiring information (examples) may vary across hypotheses. Because of these varying benefits and costs sometimes significant benefits can be derived from not bounding the statistical error equally across each of the pairwise comparisons. The STOP1 algorithm, which does not account for these varying benefits and costs, uses equal bounds across the pairwise comparisons. The STOP2 algorithm estimates the likely cost and benefit for each new example and allocates examples to the comparison with the highest estimated benefit divided by cost. This can result in a situation where each comparison is estimated to a different level of statistical error, although the sum of these errors still must remain below the overall bound of $1 - \gamma$. As the individual pairwise confidences may vary, we introduce the new notation γ^*_{ij} to signify the confidence that strategy H_i dominates or is indifferent with strategy H_j.

For example, as shown in Figure 2, if the uncertainty in determining the dominance of H_i over H_k has already been reduced significantly, and the uncertainty in showing the dominance of H_i over H_j has not, additional examples to H_i vs. H_j are likely to have greater effect on reducing the overall error than examples from H_i vs. H_k. Thus one can estimate the *marginal benefit* of allocating additional samples, the reduction in statistical error resulting from an additional example, by assuming that the mean and variance of U_{j-i} will change little and computing the increase in certainty. This results in the following formula.

$$\Delta\gamma_{ji} \text{ for } U_{j-i} = \Phi(-\overline{U}_{i-j}\frac{\sqrt{n+1}}{S_{i-j}}, \infty) - \text{ old } \gamma^*_{ji}$$

(marginal confidence dominance equation)

Similarly, we estimate the marginal increase in indifference confidence from acquiring an

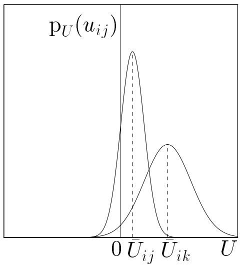

Fig. 2. Varying Effects of Sampling

additional example of U_{j-i} as follows:

$$\Delta\gamma_{ji} \text{ for } U_{j-i} = \Phi(\frac{(\overline{U}_{i-j}-\epsilon)\sqrt{n+1}}{S_{i-j}}, \frac{(\overline{U}_{i-j}+\epsilon)\sqrt{n+1}}{S_{i-j}}) - \text{ old } \gamma^*_{ji}$$

The second factor considered by STOP2 and not by STOP1 is the varying cost of acquiring a sample. If acquiring an additional sample has an extremely high cost, it may not be worth the effort, even if the expected information gain is large. Likewise, a low information cost may make a lesser information gain look more attractive. To decide how best to allocate learning resources, STOP2 estimates marginal cost. This is the cost of acquiring another sample for a given pairwise comparison and it consists of the cost of determining a utility value for each member of the pair. As each comparison shares the same hypothesis H_{High}, at least part of this cost may already have been incurred. Thus estimating the marginal cost involves two parts. First, determine which utility values must be determined (U_i, U_j, or both). Second, use the estimated means for C_i and C_j to estimate the cost of acquiring another sample U_j and U_i as appropriate.

The estimated marginal cost of determining another point from U_{j-i} is computed as follows. Let N_{ai} indicate the number of samples drawn from the strategy H_i so far. When we draw a problem from the distribution, we store it so that if we wish to sample p times from the distribution U_i, and p times from distribution U_j, we have the same p problems from the problem distribution. Furthermore, when we compute differences in utility from the distribution U_{j-i} these are computed by using the competing strategies on the same problem. Thus if we wish to get the pth sample from the distribution U_{j-i}, (assuming $p-1$ samples have already been computed), N_{ai} and N_{aj} must each be at least $p-1$. The cost can be expressed as follows:

- If both N_{ai} and N_{aj} are p or greater: the cost of computing the pth sample is 0.
- If $N_{ai} = p-1$ and $N_{aj} = p$ then the expected cost is $\overline{C}_i$.
- If $N_{aj} = p-1$ and $N_{ai} = p$ then the expected cost is $\overline{C}_j$

- If $N_{aj} = p - 1$ and $N_{ai} = p - 1$ then the expected cost is $\overline{C}_i + \overline{C}_j$

Given the marginal benefit and marginal cost, STOP2 uses the common greedy approach of selecting the course of action which has the highest ratio of marginal return to marginal cost. This process continues until a strategy emerges which can be shown with overall confidence γ to be dominant or indifferent with respect to all other strategies.[6]

```
STOP2 ALGORITHM:
let $T = H_1, \ldots, H_k$
solve $n_0$ problems with each strategy in T
compute utility comparison statistics for $n_0$ samples
LOOP1
    let $H_High$ be the strategy in T with the highest sample mean $\overline{U}$
    if for every $H_j$ in T one of the following conditions holds
        $U_{High}$ dominates $U_J$ with confidence $\gamma^*$
        $U_{High}$ and $U_J$ are ambivalent with confidence $\gamma^*$
        such that $\sum \gamma^* \leq \gamma$
    then return $H_{High}$
    else for each strategy $H_i$ in T
        Compute the marginal benefit $MB_i$ and marginal cost $MC_i$
        of acquiring another sample from $U_{High-i}$
    for the $H_i$ with the highest $MB_i/MC_i$
    generate data for the distribution $U_{High-i}$
    recompute utility comparison statistics, reselecting
    $H_{High}$ if necessary
CONTINUE WITH LOOP1
```

Again, the algorithm has been simplified to ease understanding. In fact, the marginal cost and utility of acquiring another sample need only be updated when relevant samples are taken. Additionally, acquiring a sample for H_{High} to acquire a sample for U_{High-i} may allow another U_{High-j} to be computed at zero cost (due to changes in H_{High}) and hence should be included in the relevant marginal benefit calculation.

[6] Note that in general, the system will be attempting to show that a specific strategy H_i dominates or is ambivalent with all the others. This means that N_{ai} will be consistently greater than or equal to all other N_{aj}. Anytime N_{ai} incremented to find out more information regarding H_i, this immediately reduces the cost of acquiring information for other H_j's, as they no longer need to pay the cost of sampling H_i. This will tend to mitigate the effects of different means and variances for U_{j-i} distributions. However, in cases where the focus strategy H_i changes, other more complex phenomena will occur.

IV. The Expected Loss Approach

A commonly used measure in valuing information in game theory applications is the concept of expected loss. Put simply, expected loss is the chance that one makes the wrong decision, weighted by how wrong the decision turns out to be. The expected loss measure can be computed for any pair of alternatives. These computed values can then be used to answer both the question of "is the current information enough" and if additional information is needed "which information at which cost should we get". The former question can be answered by putting a bound on the expected loss that one is willing to tolerate, and making a decision when an alternative is found to have an expected loss of less than the bound. In our case of hypothesis evaluation, one can select a hypothesis strategy H_i when: $\sum_{j=1}^{k} E[L(H_i, H_j)] \leq L$ [7] More rigorously, we define the expected loss of utility from adopting H_i rather than H_j to be the integral of the joint utility of H_i and H_j over the regions where H_i has lower utility weighted by the difference in utility:

$$E[L(H_i, H_j)] = \int_{u_i < u_j} \int P_{U_i U_j}(u_i, u_j)(u_j - u_i)\, du_i\, du_j$$

However, because U_i and U_j are jointly gaussian, and a linear combination of two jointly gaussian random variables is gaussian, we can use the differential distribution U_{i-j} to compute the expected loss directly.

Thus we simply estimate the mean and variance for our best guess at the true mean of the differential distribution U_{i-j}. [8]

We compute the integral over the region where $U_{i-j} > 0$ of the term $u \cdot Pr(U_{i-j} = u)$. To do this, we first compute the sample mean and variance for the differential distribution, and then apply a formula analogous to that used in the dominance confidence interval calculation (for derivation, see Appendix A).

$$E[L(H_i, H_j)] = \frac{S_{i-j} e^{-0.5n(\frac{\overline{U}_{i-j}}{S_{i-j}})^2}}{\sqrt{2\pi n}} + \frac{\overline{U}_{i-j}}{\sqrt{2\pi}} \int_{\frac{\overline{U}_{i-j}\sqrt{n}}{S_{i-j}}}^{\infty} e^{-0.5z^2} dz$$

A. The EL1 Algorithm

Given this definition of expected loss, we can define the analogs of STOP1 and STOP2, called EL1 and EL2.

EL1 ALGORITHM:

```
let $T = H_1, \ldots, H_k$ and L be the expected loss threshold.
let L* = L/k
```

[7] Enforcing that E(L(H,H))=0.

[8] An alternative approach would be to estimate the parameters for each of the individual utility distributions, then use these parameters to compute the mean and variances for the estimates of the differential distributions. This would result in the same parameters as our approach of computing the parameters of the differential distributions directly from the data.

```
solve $n_0$ problems with each strategy in T
let $H_{High}$ be the strategy in T with the highest sample $\overline{U}$
$\forall j = 1, \ldots, k$ compute $E(L(H_{high}, H_j))$
LOOP1
    select a strategy $H_i$ such that the
            expected utility loss from selecting
            $H_{High}$ over $H_i$ is greater than L*
    if there is no such strategy,
    then return $H_{High}$
    else generate sample from $H_i$ and $H_{High}$
            recompute expected utility losses
CONTINUE LOOP1
```

B. The EL2 Algorithm

EL2 extends EL1 in exactly the same way that STOP2 extends STOP1, by accounting for variable gains and costs across the hypotheses.

The *marginal decrease in expected utility loss* (MDEUL) is computed by recomputing the integral for expected loss, assuming that the variances and means will remain the same but incrementing n by 1 and subtracting the current expected utility loss. The resulting formula is shown below.

$$\begin{aligned} \Delta E[L(H_i, H_j)] &= \frac{S_{u_i - u_j} e^{-0.5n(\frac{\overline{U}_{i-j}}{S_{i-j}})^2}}{\sqrt{2\pi(n+1)}} \\ &+ \frac{\overline{U}_{i-j}}{\sqrt{2\pi}} \int_{\frac{\overline{U}_{i-j}\sqrt{n+1}}{S_{i-j}}}^{\infty} e^{-0.5z^2} dz - \text{ old } E[L(H_i, H_j)] \end{aligned}$$

The expected marginal cost of sampling is computed as in STOP2. The EL2 algorithm is shown below.

```
EL2 ALGORITHM
let $T = H_1, \ldots, H_k$ and L be the expected loss threshold.
solve $n_0$ problems with each strategy in T
let $H_{High}$ be the strategy in T with the highest sample mean $\overline{U}$
$\forall i = 1, \ldots, k$ compute $E[L(H_{High}, H_i)]$
    and let this be $l_i^*$ (enforce that $E(L(H_i, H_i)) = 0$
loop1
    if $\sum l^* \leq L$
    then return $H_{High}$
    else compute the marginal decrease in expected
        loss (MDEUL) by sampling from each of the
        $H_i$'s (including $H_{High}$)
        compute the marginal cost of sampling each
        strategy using the C distributions
        sample from the distribution with the highest
        MDEUL/expected marginal sampling cost
        recompute $L_i^*$'s as necessary
    continue loop1
```

V. Empirical Performance Evaluation

We now turn to an empirical evaluation of the hypothesis selection techniques. This evaluation lends support to the techniques by addressing three key issues. First it demonstrates that the techniques perform as predicted. Second, the evaluation demonstrates the benefits of rational example allocation (as performed by STOP2 and EL2). Finally, it illustrates the applicability of the approach to a real-world hypothesis selection problem. Where possible, we contrast performance with that of other relevant approaches in the statistical literature.

A. Other Relevant Approaches

There exists a body of standard approaches for the interval-based formulation of the hypothesis evaluation problem. To demonstrate the power of our interval-based approaches we contrast them with two existing approaches. The first is a statistical approach proposed by Turnbull and Weiss [21]. The second is the COMPOSER machine learning technique proposed by Gratch and DeJong [1].

The Turnbull and Weiss approach comes closest among statistical ranking and selection procedures to the generality of the STOP1 and STOP2 approaches. Most standard statistical approaches make strong assumptions about the form of the hypothesis evaluation problem (e.g., the variances associated with hypotheses are known or equal). As in our interval-based approaches, Turnbull and Weiss treat hypotheses as normal variables with unknown mean, and unknown and unequal variance, however they make the additional assumption that hypotheses are independent. It can still be reasonable to use this approach when the hypotheses are not independent, but this can lead to excessive statistical error or unnecessarily large training set sizes under certain circumstances. However, in the case where hypotheses are truly independent, this technique can exploit this knowledge and likely outperform our methods which do not adopt this assumption. The Turnbull technique is described in Appendix B.

The COMPOSER technique (described in Appendix C) was proposed to solve hypothesis evaluation problems as they arise in the context of adaptive problem solving. COMPOSER treats hypotheses as dependent normal variables with unknown mean, and unknown and unequal variance. COMPOSER, however, does not implement the notion of an indifference interval. Rather it is trying to adopt the first hypothesis that can be demonstrated to be significantly better than a default hypothesis. When the best hypotheses are all close to each other in utility, COMPOSER will require an excessive number of training examples.

B. Methodology

First we discuss some methodological issues. The interval-based and expected loss approaches embody different criteria for selecting hypotheses and therefore are difficult to compare directly. Thus we first test the interval-based and expected loss approaches separately. Interval-based approaches have been investigated extensively in the statistical ranking and selection literature (see [22] for a review of the recent literature). This affords us the opportunity to compare STOP1 and STOP2 against a standard statistical approach.

Techniques are evaluated on synthetic and real-world data sets. Synthetic data allows a systematic test of the formal properties of each technique while real data sets test the appropriateness of statistical assumptions – such as the normal approximation – and assess the practicality of each approach on real-world problems. Finally, in a comprehensive real-world test on scheduling data, we compare the interval based and expected loss approaches, using a wide range of parameter settings. This test reports on the bottom-line effectiveness of the competing techniques in a pragmatic problem-solving setting.

An experimental trial consists of solving a hypothesis evaluation problem with a given technique. The performance on any single trial provides little information given the random nature of the task. To assess the average characteristics of the technique a trial is repeated multiple times and the results are averaged across trials. All experimental trials are repeated 5000 times.

An interval-based technique processes examples until it has identified a hypothesis that with probability γ^* is within ϵ of optimal. STOP1 attempts to ensure this property with the minimum number of training examples possible. STOP2 attempts to ensure this property with the minimum cost possible. To assess the competence of these techniques we track three quantities: the number of examples required to choose a hypothesis, the cost of the examples required to choose a hypothesis, and the observed probability that the expected utility of the chosen hypothesis is in fact within ϵ of the utility of the optimal hypothesis. For the expected loss techniques we track the analogous three quantities: the number of examples to choose a hypothesis, the cost of the examples, and the average loss (the average loss in utility when the technique chooses the nonoptimal hypothesis weighted by the probability of choosing the nonoptimal hypothesis).

B.1 Synthetic Data

Synthetic data is used to show that: (1) the techniques perform as expected when the underlying assumptions are valid and (2) the use of rational example allocation exhibits substantial improvement when there is unequal cost or variance among the distributions. For interval based approaches we show that the technique will choose the best hypotheses, or one ϵ-close to the best, with the requested probability. When all hypotheses are within ϵ of each other, the indifference-based technique should quickly terminate, returning any hypotheses. For the expected loss approaches the claim is that the technique will exhibit no more that the requested level of expected loss. One set of evaluations is devised to test this claim.

The second claim is that the techniques that use rational example allocation will exhibit substantial performance improvement when there is unequal cost or variance among the hypotheses. A second set of evaluations is devised to test this claim

For the synthetic data problems, hypotheses are modeled as random variables with parameterized properties. A specific hypothesis evaluation problem is constructed by fixing the values of each of these parameters. In the course of solving a specific problem, values for the utility and cost of each hypothesis on each example are assigned randomly according to the parameterized distribution functions. For a given problem let k define the number of hypotheses. For all synthetic evaluations, the hypothesis utilities and costs are treated as independent normal random variables with some parameterized mean and variance. Each hypothesis is described by four parameters – expected utility, utility variance, expected cost, and cost variance. Thus a hypothesis evaluation problem is specified by 4k parameters.

The hypothesis evaluation techniques have additional parameters that govern how they attack the problem. To distinguish these we refer to problem parameters and control parameters. The interval-based techniques have three control parameters: an initial sample size n_0, a confidence setting γ^* and an indifference setting ϵ. The expected loss techniques have two control parameters: an initial sample size n_0 and a loss threshold H*.

Unless otherwise stated, each training example on any hypothesis is given equal cost. This means that the overall cost of a technique is directly proportional to the expected number of examples required to select a hypothesis. Thus, when each training example is given equal cost only the number of examples will be reported. One set of synthetic evaluations highlights the benefits of rational example allocation. In these evaluations we create a significant discrepancy in the cost of evaluating alternative hypotheses.

B.2 Scheduling Data

The test of real-world applicability is based on data drawn from an actual NASA scheduling application [23]. This data provides a strong test of the applicability of the techniques. All of the statistical techniques make some form of normality assumption. However the data in this application is highly non-normal – in fact most of the distributions are bimodal. This characteristic provides a rather severe test of the robustness of the approaches.

In this application a heuristic system was developed to schedule communication events between earth-orbiting satellites and ground based radio antennas. In the course of development, extensive evaluations were performed with various scheduling heuristics. The goal of these evaluations was to choose a heuristic search strategy that solved scheduling problems quickly on average. This is easily seen as a hypothesis evaluation problem. Each of the heuristics corresponds to a hypothesis. The cost of evaluating a hypothesis over a training example is the cost of solving the scheduling problem with the given heuristic. The utility of the training example is simply the negation of its cost. In that way, choosing a hypothesis with maximal expected utility corresponds to choosing a scheduling heuristic with minimal average cost.

Using the data from the heuristic evaluations we derived four data sets. Each data set corresponds to a comparison of some set of scheduling heuristics, and contains data on the heuristics' performance over about one thousand scheduling problems. An experimental trial consists of executing a technique over one of these data sets. Each time a training example is to be processed, some problem is drawn randomly from the data set with replacement. The actual utility and cost values associated with this scheduling problem is then used. As in the synthetic data, each experimental trial is repeated 5000 times and all reported results are the average of these trials.

C. The Interval-Based Approach

The interval-based approaches, STOP1 and STOP2, are evaluated on both synthetic and scheduling data sets. Synthetic problems were constructed to answer the following three questions: 1) do the techniques select ϵ-close hypotheses with the specified probability, 2) do the techniques terminate quickly when all hypotheses are ϵ -close, and 3) does STOP2 outperform STOP1 when there is significant cost or variance differences between hypotheses. We also contrast the performance of our techniques with COMPOSER and the technique of Turnbull and Weiss.

C.1 Confidence Test

The statistical ranking and selection literature uses a standard methodology for evaluating the statistical error of hypothesis evaluation techniques. We adopt this methodology here. Robert Bechhofer introduced the concept of the least favorable configuration of the population means [18]. This is a parameter configuration that is most likely to cause a technique to choose a wrong hypothesis (one that is not ϵ-close) and thus provides the most severe test of the technique's abilities. Under this configuration, k - 1 of the hypotheses have identical expected utilities, μ , and the remaining hypothesis has expected utility $\mu + \epsilon$. The last hypothesis has the highest expected utility and should be chosen by the technique. All hypotheses are independent and the costs and variances of all hypotheses are equal. [9]

The least favorable configuration becomes more difficult (requires more examples) as the

[9] Note that in this evaluation ϵ acts as a problem parameter in addition to its role as a control parameter.

TABLE I

ESTIMATED EXPECTED TOTAL NUMBER OF OBSERVATIONS IN THE LEAST FAVORABLE CONFIGURATION. ACHIEVED PROBABILITY OF CORRECT SELECTION IS SHOWN IN PARENTHESIS.

k	γ^*	$\frac{\sigma}{\epsilon}$	STOP1	STOP2	TURNBULL	COMPOSER
3	0.75	2	38 (0.85)	34 (0.83)	27 (0.75)	61 (0.96)
3	0.75	3	58 (0.08)	52 (0.78)	50 (0.72)	103 (0.90)
3	0.90	2	64 (0.92)	65 (0.92)	54 (0.86)	91 (0.98)
3	0.90	3	121 (0.91)	123 (0.91)	127 (0.87)	170 (0.95)
3	0.95	2	93 (0.95)	96 (0.97)	81 (0.92)	115 (0.99)
3	0.95	3	183 (0.94)	193 (0.95)	192 (0.93)	238 (0.97)
5	0.75	2	98 (0.86)	94 (0.86)	63 (0.71)	139 (0.96)
5	0.75	3	177 (0.83)	179 (0.81)	141 (0.71)	250 (0.89)
5	0.90	2	159 (0.93)	170 (0.94)	123 (0.84)	195 (0.97)
5	0.90	3	310 (0.92)	349 (0.93)	294 (0.88)	389 (0.94)
5	0.95	2	212 (0.96)	234 (0.97)	175 (0.91)	237 (0.98)
5	0.95	3	427 (0.95)	483 (0.96)	411 (0.94)	501 (0.97)
10	0.75	2	298 (0.89)	330 (0.90)	185 (0.66)	353 (0.95)
10	0.75	3	584 (0.87)	688 (0.87)	438 (0.70)	677 (0.89)
10	0.90	2	430 (0.95)	508 (0.95)	331 (0.83)	469 (0.97)
10	0.90	3	892 (0.93)	1,066 (0.95)	783 (0.89)	958 (0.93)
10	0.95	2	545 (0.97)	661 (0.97)	443 (0.91)	574 (0.98)
10	0.95	3	1,136 (0.95)	1,435 (0.97)	1,037 (0.94)	1,175 (0.95)

confidence γ^*, the number of hypotheses k, or the common utility variance σ^2 increases. It becomes easier as the indifference interval ϵ increases. In the standard methodology a technique is evaluated using several settings for k , γ^*, and $\frac{\sigma}{\epsilon}$. The last term combines the variance and indifference interval size into a single quantity which, as it increases, makes the problem more difficult. For our experiments, $n_0 = 7$, $\mu = 50$, $\sigma = 64$, and all other parameters are varied as indicated in the results. The sample size results and observed confidence levels are summarized in Table I.

The results indicate that all systems are roughly comparable in the number of examples required to choose a hypotheses. As expected, the number of examples increases with k, γ^*, and $\frac{\sigma}{\epsilon}$. The technique of Turnbull and Weiss tended to be the most efficient, however this algorithm was essentially told that the hypotheses are independent; information that was withheld from the other algorithms. COMPOSER performed the worst of the algorithms. In terms of statistical error, all of the algorithms except Turnbull and Weiss' were correct

TABLE II

Estimated expected total number of observations in the idifference configuration. Note that COMPOSER failed to terminate on any of the trials.

Parameters			STOP1	STOP2	TURNBULL	COMPOSER
k	γ^*	$\frac{\sigma}{\epsilon}$				
3	0.75	2	48	44	27	***
3	0.75	3	75	68	50	***
3	0.90	2	96	100	54	***
3	0.90	3	181	194	127	***
3	0.95	2	142	151	81	***
3	0.95	3	291	312	192	***
5	0.75	2	134	143	63	***
5	0.75	3	249	276	141	***
5	0.90	2	235	267	123	***
5	0.90	3	474	568	294	***
5	0.95	2	325	360	174	***
5	0.95	3	672	768	411	***
10	0.75	2	421	525	185	***
10	0.75	3	833	1104	438	***
10	0.90	2	649	772	331	***
10	0.90	3	1348	1667	782	***
10	0.95	2	835	975	444	***
10	0.95	3	1776	2100	1037	***

at least as often as requested. The technique of Turnbull and Weiss often provided less than the requested confidence. However, since their technique only guarantees that the confidence will approach γ^* as $\frac{\epsilon}{\sigma}$ tends to zero, these results are consistent with their claim.

C.2 Indifference Test

The indifference interval approaches should terminate quickly when all hypotheses are indifferent to each other. To test this claim we repeated the least favorable configuration evaluations except that all hypotheses were assigned the same expected utility μ. Results are summarized in Table II. Error rate results are not shown since any hypothesis is a correct selection in this configuration.

The key result to notice is that COMPOSER failed to terminate on any of the trials. This highlights the potential difficulties with COMPOSER that STOP1 and STOP2 were designed to correct. Again, the technique of Turnbull and Weiss could exploit the

TABLE III

Estimated expected total cost for the rational allocation configuration

k	γ^*	STOP1	STOP2	$\frac{STOP1}{STOP2}$
3	0.75	12,034	5,241	2.3
3	0.80	14,890	6,790	2.2
3	0.85	20,119	10,030	2.0
3	0.90	26,340	15,040	1.8
5	0.75	22,081	5,216	4.2
5	0.80	27,375	6,947	3.9
5	0.85	31,203	9,817	3.2
5	0.90	39,305	14,859	2.7
10	0.75	36,768	5,154	7.1
10	0.80	42,202	6,753	6.3
10	0.85	47,167	10,086	4.7
10	0.90	56,183	15,004	3.8

independence information and slightly outperforms the other approaches.

C.3 Rational Allocation Test

STOP2 is designed to perform well when the cost of processing examples or the utility variance differs widely across hypotheses. The preceding evaluations did not contrast the two approaches under these conditions as both the cost and variances were equal. Consequently STOP1 and STOP2 were approximately equally efficient in these tests. This evaluation contrasts the approaches by providing problem configurations with highly unequal costs.

Problem configurations are defined as follows. One hypothesis (the correct selection) is assigned a high mean μ_{best}. A second hypothesis is assigned a mean slightly below ϵ of the best, μ_{best-1}. All remaining hypotheses are assigned a low mean, μ_{worst}. The second hypothesis is given a high cost c_{high} and all other hypotheses are given low cost c_{low}. All hypotheses are assigned a common variance of fifty, $\mu_{best} = 74$, $\mu_{best-1} = 72$, $\mu_{worst} = 5$, $\epsilon = 1$, and $n_0 = 7$. Various confidence settings were evaluated. The results are summarized in Table III.

The results illustrate the clear dominance of STOP2 under this configuration - up to seven times more efficient on one of the trials. An interesting question is whether there is a limit to how much better STOP2 can be. In fact there is an upper bound on this

TABLE IV

Estimated expected total number of observations for scheduling data. Achieved probability of a correct selection is shown in parenthesis.

set	k	γ^*	$\frac{\sigma}{\epsilon}$	STOP1	STOP2	TURNBULL	COMPOSER
D1	3	0.95	34	908 (1.00)	648 (1.00)	26,691 (1.00)	78 (1.00)
D2	2	0.95	34	74 (1.00)	76 (1.00)	13,066 (1.00)	346 (1.00)
D3	7	0.95	14	2,371 (0.94)	2,153 (0.93)	94,308 (1.00)	2,456 (0.97)
D4	7	0.95	11	7,972 (0.96)	7,621 (0.94)	87,357 (1.00)	21,312 (0.89)

difference [24]. This upper bound increases as the number of hypotheses increases or as the confidence level decreases.

C.4 Scheduling Test

We ran all four algorithms over the four scheduling data sets. In each case the $\gamma = 95\%, n_0 = 15$, and $\epsilon = 4.0$. Table IV summarizes the results along with the number of hypotheses and the relative difficulty ($\frac{\sigma}{\epsilon}$) of each data set.

The principle result is that STOP1 and STOP2 substantially exceeded the performance of the other algorithms except on one case. The one exception is an artifact of COMPOSER solving a slightly different task. Rather than choosing the hypothesis that is ϵ-close to optimal, COMPOSER chooses the first hypothesis to dominate a default hypothesis (the first hypothesis was arbitrarily defined to be the default in these trials). In data set D1 the default is significantly worse than the other two hypotheses, which in turn are indifferent to each other. STOP1 and STOP2 take longer because they must verify this indifference.

Note that unlike the synthetic data where STOP1 was slightly more efficient than STOP2, in the scheduling data STOP2 was slightly more efficient. In fact, in the scheduling data there is some disparity between hypotheses in their utility variance. STOP2 is able to account for these factors when allocating examples, and thus exhibits greater efficiency.

Turnbull and Weiss' technique performed substantially worse on the real-world data. Its poor performance is due to two factors. First, the technique is unable to quickly discard hypotheses that are clearly dominated by other hypotheses. Second, the technique's independence assumption was inappropriate for this data, which is strongly positively correlated. In this situation assuming independences leads to overestimates of the true variance, which in turn leads to higher sample sizes.

D. Discussion of Interval-Based Evaluation

Taken together, the evaluation provides clear evidence for the effectiveness of STOP1 and STOP2 and demonstrates their superiority to alternative techniques. The techniques performed as predicted, guaranteeing the requested confidence level under a variety of configurations. In comparison to other approaches, they did not perform the best on every configuration, however when they were outperformed it was not by much and they often substantially outperformed the alternative techniques. For example, COMPOSER fails to terminate when multiple hypotheses are close to optimal. The technique of Turnbull and Weiss performed poorly on the real-world data sets. The scheduling evaluation demonstrates that STOP1 and STOP2's normal approximation allows effective performance on real-world hypotheses selection problems, even when the underlying distributions are not normal.

The rational allocation test illustrates that STOP2 can substantially outperform STOP1 when there are marked differences across heuristics in the cost of processing examples or in the variance of expected utility values. STOP2 should be used if the hypothesis evaluation problem has this characteristic. It appears that STOP1 is slightly more efficient when the cost and utilities are close to equal. Under these circumstances we recommend the use of STOP1.

E. The Expected Loss Approach

The expected loss approaches, EL1 and EL2, are evaluated on both synthetic and scheduling data sets. Synthetic problems are constructed to answer the following two questions: 1) do the techniques properly bound the expected loss, and 2) does EL2 outperform EL1 when there is significant cost or variance differences between hypotheses.

E.1 Expected Loss Test

The techniques are tested on a least favorable configuration with k hypotheses. The means of $k-1$ hypotheses are assigned the value m and the remaining hypothesis is

TABLE V

Estimated expected total number of observations and expected loss of an incorrect selection for the least favorable configuration.

Parameters			EL1		EL2	
k	ϵ	H^*	Samples	Loss	Samples	Loss
3	2	1.0	33	0.5	26	0.8
3	2	0.75	38	0.4	29	0.7
3	2	0.5	46	0.2	35	0.5
3	2	0.25	58	0.1	48	0.3
5	2	1.0	73	0.4	54	0.9
5	2	0.75	83	0.3	62	0.7
5	2	0.5	98	0.2	78	0.5
5	2	0.25	127	0.1	114	0.2
10	2	1.0	201	0.2	157	0.8
10	2	0.75	221	0.2	182	0.6
10	2	0.5	255	0.1	220	0.4
10	2	0.25	312	0.0	269	0.2

assigned mean $m + \epsilon$. Each technique is then tested on various loss thresholds H* over this problem. For this evaluation, $m = 50$, all hypotheses share a common utility variance $\sigma^2 = 64$, and $\epsilon = 2$. All other parameters are varied as indicated in the results. The sample size results and observed loss values are summarized in Table V. The results illustrate that the techniques perform as predicted. As the loss threshold is lowered the techniques take more training examples to ensure the expected loss remains below the threshold.

E.2 Rational Allocation Test

EL2 is designed to perform well when the cost of processing examples or the utility variance differs widely across hypotheses. The preceding evaluations did not contrast the two techniques as the cost and variances were equal across hypotheses. This evaluation contrasts the approaches using unequal costs across the hypotheses. The configuration used is identical to the one described in Section C.3. The difference in expected costs between solving problems with EL1 and EL2 is summarized in Table VI. The results indicate that EL2 substantially outperformed EL1 – in one trial solving the configuration four times more efficiently. EL2 achieves greater efficiency as the number of hypotheses increases. As with STOP2 we suspect that the potential for greater efficiency is not

TABLE VI

Estimated expected total cost for the rational allocation configuration.

k	H^*	EL1	EL2	$\frac{EL1}{EL2}$
3	1.00	5,757	3,733	1.5
3	0.75	6,980	3,992	1.8
3	0.50	8,899	4,636	1.9
3	0.25	14,102	6,847	2.1
5	1.00	8,070	3,737	2.2
5	0.75	9,688	3,985	2.5
5	0.50	12,807	4,664	2.8
5	0.25	19,525	6,873	2.9
10	1.00	12,745	3,740	3.2
10	0.75	15,035	4,037	3.7
10	0.50	19,144	4,718	4.1
10	0.25	26,901	6,861	3.9

TABLE VII

Estimated expected total number of observations and expected loss of an incorrect selection for the scheduling data.

	Parameters		EL1		EL2	
set	k	H^*	Samples	Loss	Samples	Loss
D1	3	3.0	78	0.1	49	1.0
D2	2	3.0	30	1.8	30	1.8
D3	7	3.0	335	3.0	177	3.9
D4	7	3.0	735	1.7	283	2.2

unbounded, but we have not as yet obtained an upper bound on the relative efficiency of EL2.

E.3 Scheduling Test

We ran the two expected-loss based techniques over the four scheduling data sets. In each case the L=3 and $n_0 = 15$. The results are shown below in Table VII. The main result is that the algorithms correctly bounded the expected loss with one exception – EL2 gave greater than expected loss on data set D3. It appears that this exception arose from a significant departure from normality in the distributions comprising the data set. Additional trials demonstrated this discrepancy goes away if the initial sample size is increased, thereby improving the normal approximation.

F. Discussion of Expected Loss Evaluation

The three evaluations of EL1 and EL2 give clear support for the effectiveness of these algorithms. The techniques performed as predicted, properly bounding the expected loss under a variety of parameter configurations. We did observe that under some of the configurations EL2 gave slightly larger than requested loss. More generally, it appears that the expected loss approach will be more susceptible to departures from normality in the utility distributions, when compared with interval-based approach. Both approaches use a normal distribution to approximate the distribution of a sample mean. However the interval-based approach is only sensitive to the area under parts of the normal curve. The expected loss computation makes use of both the area and the shape of certain parts of the normal curve. Thus the expected loss approach demands more fidelity from its approximation, and this fidelity is degraded when the underlying distribution is not normal. This effect can be compensated by using a larger n_0 for the expected loss technique.

G. Comparing Interval-based to Expected Loss

One cannot state that interval-base techniques are better or worse than expected loss approaches – each is solving a slightly different problem. Interval-based approaches are attempting to identify a nearly optimal hypothesis with high confidence while expected loss approaches are attempting to minimize the cost of a mistaken selection. If the goal of the task is to identify the best hypothesis then clearly an interval-based approach should be used. If the goal is to simply improve expected utility as much as possible, either could be used and it is unclear which is to be preferred.

Our original motivation in developing these approaches was to develop effective techniques for adaptive problem solving. In this section we attempt to assess how the various approaches perform on this task. In particular we consider how the approaches perform in the problem of learning a set of problem solving heuristics for the NASA scheduling domain. In this test the algorithms were given the task of optimizing four control parameters of the adaptive scheduler, with the goal of speeding up the schedule generation process. The solution to this consists of identifying a good heuristic for each of the four control parameters, where the best choice for a particular parameter depends on the heuristics

TABLE VIII

Direct comparison of all four algorithms.

Algorithm	Cost (100s CPU sec)	Examples	Utility
COMPOSER (0.90)	6128	4075	17.3
STOP1 (0.75,1.0)	4199	2785	17.1
STOP2 (0.75,1.0)	3140	1924	16.6
EL1 (1.0)	2347	1557	16.8
EL2 (0.5)	2211	1454	16.4

chosen for the other control parameters. We implement a hillclimbing strategy for finding a good combination of heuristics. For more details on this application domain see [23].

We run each algorithm under a variety of parameter settings and compare the best performance of each algorithm (i.e., the lowest cost setting that resulted in a high expected utility on average). In this test the interval-based algorithms are run with confidence levels γ^*=0.75,0.90,0.95 and indifference levels ϵ=1.0, 4.0, 7.0. The expected loss algorithms are run with loss bound L=5, 1, 0.5. For each setting 1000 runs are conducted, we then determined the best settings as the lowest cost solution within 1.0 utility of the average best solution found per algorithm (effectively enforcing a minimum utility of 16.5). These best settings and the averaged results (from 1000 runs each) are shown in Table VIII. These results show that the algorithms produce roughly comparable utilities, the difference in utilities is smaller than the smallest indifference interval specified to the interval-based algorithms. From this comparison we must conclude that, at least in the case of this NASA scheduling application, there is little difference between the interval-based and expected loss approaches, neither in terms of expected improvement nor in terms of sample complexity. As expected, the unequal allocation approaches performed better in terms of learning cost. Finally, all of the improved algorithms outperformed the benchmark COMPOSER algorithm in terms of learning cost.

VI. Discussion and Conclusions

There are many issues relevant to hypothesis evaluation which have not been addressed in this paper. One issue is modeling the computational cost of inferring and applying the statistical models. In some applications, one might imagine that these costs would play a significant role in determining the usefulness of our hypothesis evaluation mode. However, in our target application of learning for scheduling, the cost of gathering further information heavily outweighs the cost of inferring and applying the statistical models. However, for other domains we concede that this may not be the case. A second related issue is to estimate and tradeoff this cost of applying the statistics and decision theory relative to the cost of additional examples.

Another issue is to better understand the qualitative conditions under which the cost sensitive measures (STOP2 and EL2) will outperform the equal error distribution models (STOP1 and EL1). Generally speaking, if the means and variances vary significantly, the cost sensitive measures should perform better. Additionally, if the marginal computations are reasonable projections, the cost sensitive measures should also outperform the other measures.

An important issue is the use of the $O(k)$ error function. Further empirical evaluation needs to be performed to better understand the relationship between n_0 and the number of H_{high} switches during hypothesis evaluation, and exactly how this relates to the error models and to the required confidence parameter γ. As a further subtlety, one might consider removing strategies which become dominated at any point in the evaluation (in contrast with the current approach which requires all strategies to be compared against the final H_{high}).

Another issue is determining the exact impact of the dual example phenomenon (where two examples are needed to compute each data point for the differential distribution) would be desirable. Additionally, if we had a method of estimating a utility difference with unequal numbers of examples that would be very helpful, but since the utilities are covarying it seems unlikely that such a technique will be found.

This paper has described techniques for choosing among a set of alternatives in the presence of incomplete information and varying costs of acquiring information. In our approach, the cost and utility of various alternatives are represented using parameterized statistical models. Using techniques from an area of statistics called parameter estimation, models can be inferred from performance on sample problems. These statistical models can then be used to estimate the utility and cost of acquiring additional information and the utility of selecting specific alternatives from the possible choices at hand. These techniques have been applied to adaptive problem-solving, a technique in which a system automatically tunes various control parameters on a performance element to improve performance in a given domain. Empirical results were presented comparing the effectiveness of these techniques on artificially generated data and speedup learning from a real-world NASA scheduling domain.

Acknowledgements

Portions of this work were performed by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration and portions at the Beckman Institute, University of Illinois under National Science Foundation Grant NSF-IRI-92-09394

References

[1] J. Gratch and G. DeJong, "COMPOSER: A Probabilistic Solution to the Utility Problem in Speedup Learning," in *Proceedings of the National Conference on Artificial Intelligence*, 1992, pp. 235-240.

[2] R. Greiner and I. Jurisca, "A Statistical Approach to Solving the EBL Utility Problem" in *Proceedings of the National Conference on Artificial Intelligence*, 1992, pp. 241-248.

[3] U. Fayyad and K. Irani, "The Attribute Selection Problem in Decision Tree Generation" in *Proceedings of the Tenth National Conference on Artificial Intelligence*, 1990 pp. 104-110.

[4] R. Musick, J. Catlett, and S. Russell, "Decision Theoretic Subsampling for Induction on Large Databases" in *Proceedings of the Tenth International Conference on Machine Learning*, 1993, 212-219

[5] A. Moore and M. Lee, "Efficient Algorithms for Minimizing Cross Validation Error" in *Proceedings of the Eleventh International Conference on Machine Learning*, 1994, 190-198.

[6] L. G. Valiant, "A Theory of the Learnable," Communications of the Association for Computing Machinery 27, (1984), pp. 1134-1142.

[7] S. Russell and E. Wefald, "On Optimal Game Tree Search using Rational Meta-Reasoning," in *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence* , 1989, pp. 334-340.

[8] S. Russell and E. Wefald, *Do the Right Thing: Studies in Limited Rationality*, Cambridge, MA: MIT Press, 1991.

[9] M. Iwamoto, "A Planner with Quality Goal and Its Speed-up Learning for Optimization Problem", in *Proceedings of the Second International Conference on AI Planning Systems* 1994, pp. 281-286.

[10] S. Chien, J. Gratch, "Producing Satisficing Solutions to Scheduling Problems: An Iterative Constraint Relaxation Approach," in *Proceedings of the Second International Conference on AI Planning Systems*, 1994, pp. 213-218.

[11] M. A. Perez, J. Carbonell, "Control Knowledge to Improve Plan Quality" in *Proceedings of the Second International Conference on AI Planning Systems*, 1994, pp. 323-328.

[12] T. J. Santner and A. C. Tamhane, "Designing Experiments for Selecting a Normal Population with a Large Mean and a Small Variance" in *Design of experiments: ranking and selection*, (eds) T. J. Santner and A. C. Tamhane, Marcel Dekker, 1984.

[13] R. V. Hogg and A. T. Craig, *Introduction to Mathematical Statistics* , New York, USA: Macmillan Publishing, 1978

[14] E. Kreysig, *Introductory Mathematical Statistics: Principles and Methods*, New York, USA: John Wiley and Sons, 1970.

[15] H. Buringer, H. Martin, and K. Schriever, *Nonparametric Sequential Selection Procedures*, Boston, MA: Birkhauser, 1980.

[16] W. Yang and B. Nelson, "Using common random numbers and control variates in multiple-comparison procedures" *Operations Research* 39 4, 1991, pp. 583-591.

[17] O. Maron and A. Moore, "Hoeffding Races: Accelerating Model Selection Search for Classification and Function Approximation" in *Advances in Neural Information Processing Systems 6* , Morgan Kaufmann, 1994.

[18] R.E. Bechhofer, "A Single-sample Multiple Decision Procedure for Ranking Means of Normal Populations with Known Variances," *Annals of Math. Statistics* (25) 1, 1954 pp. 16-39.

[19] E. Paulson, "A Sequential Procedure for Selecting the Population with the Largest Mean from k Normal Populations" it Annals of Mathematical Statistics 35, 1964, pp. 174-180.

[20] J. Gratch, "COMPOSER: A Decision-theoretic Approach to Adaptive Problem-solving" Technical Report UIUCDCS-R-93-1806, Dept. of Computer Science, University of Illinois, Urbana, IL, May 1993.

[21] Turnbull and Weiss, "A class of sequential procedures for k-sample problems concerning normal means with unknown unequal variances," in *Design of Experiments: ranking and selection* , T. J. Santner and A. C. Tamhane (eds.), Marcel Dekker, 1984.

[22] R. M. Haseeb, *Modern Statistical Selection*, Columbus, OH: Am. Sciences Press, 1985.

[23] J. Gratch, S. Chien, and G. DeJong, Learning "Search Control Knowledge for Deep Space Network Scheduling," in *Proceedings of the Tenth International Conference on Machine Learning*, 1993,pp. 135-142.

[24] J. Gratch, S. Chien, and G. DeJong "Improving Learning Performance Through Rational Resource Allocation" in it Proceedings of the Twelfth National Conference on Artificial Intelligence , 1994, pp. 576-581.

Appendix A: The Expected Loss Calculation

We begin by noting that we want to integrate over the difference between the two utilities, over the region in which the unselected hypothesis strategy has a higher utility. Consider the expected loss for the selection of hypothesis strategy H_j over H_i. In order to compute this, we need to examine the differential distribution U_{i-j}, and integrate from zero to infinity.

$$E[L(H_i, H_j)] = \frac{1}{S_{i-j}\sqrt{2\pi}} \int_0^\infty e^{-0.5\left(\frac{(\overline{U}_{i-j}-l)\sqrt{n}}{S_{i-j}}\right)^2} l\, dl$$

we then make the substitution of $z = \frac{(l-\overline{U}_{i-j})\sqrt{n}}{S_{i-j}}$ which results in the following implied substitutions: $l = \frac{S_{i-j}\cdot z}{\sqrt{n}} + \overline{U}_{i-j}, dz = \frac{\sqrt{n}}{S_{i-j}} dl$ and $dl = \frac{S_{i-j}}{\sqrt{n}}\, dz$ and to compute the limits of integration we note that when $l = 0, z = -\frac{\overline{U}_{i-j}\sqrt{n}}{S_{i-j}}$ and when $l = \infty$ then $z = \frac{(\infty-\overline{U}_{i-j})\sqrt{n}}{S_{i-j}} = \infty$ resulting in:

$$E(L(H_i, H_j)) = \frac{1}{S_{i-j}\sqrt{2\pi}} \int_{-\frac{\overline{U}_{i-j}\sqrt{n}}{S_{i-j}}}^{\infty} e^{-0.5z^2} (\frac{zS_{i-j}}{\sqrt{n}} + \overline{U}_{i-j}) S_{i-j} dz$$

$$= \frac{S_{i-j}}{\sqrt{2\pi n}}(\int_{-\frac{\overline{U}_{i-j}\sqrt{n}}{S_{i-j}}}^{\infty} e^{-0.5z^2} zdz) + \frac{\overline{U}_{i-j}}{\sqrt{2\pi}}(\int_{-\frac{\overline{U}_{i-j}\sqrt{n}}{S_{i-j}}}^{\infty} e^{-0.5z^2} dz)$$

we now note that the first integral has an analytic solution, that $\int e^{-0.5x^2} xdx = e^{-0.5x^2}$ leaving us with the following:

$$E[L] = \frac{S_{i-j}e^{(\frac{\overline{U}_{i-j}}{S_{i-j}})^2}}{\sqrt{2\pi n}} + \frac{\overline{U}_{i-j}}{\sqrt{2\pi}} \int_{-\frac{\overline{U}_{i-j}\sqrt{n}}{S_{i-j}}}^{\infty} e^{-0.5z^2} dz$$

(expected loss formula 1)

Appendix B: The Turnbull and Weiss Algorithm

Turnbull and Weiss have proposed a sequential interval-based procedure for selecting the member of a population with largest mean. Members are considered normal variables with unknown mean and unknown variance. The procedure is as follows. For each hypotheses take an initial sample of n_0 observations, then take observations sequentially. Stop sampling from a hypothesis when: $\frac{S_i^2}{n_i} \leq \frac{1}{n^*}$ Where S_i^2 is the sample variance and n_i is the number of examples taken for hypothesis i. The value n^* will be defined momentarily. When sampling has stopped on all hypotheses, select the hypothesis with the highest sample mean. The value n^* is defined as $\frac{d^2}{\epsilon^2}$ where d is chosen to satisfy: $\int_{-\infty}^{\infty}(F(y+d))^{k-1}f(y)dy = \gamma^*$ where F(y) and f(y) are the cumulative distribution function and probability density function of the standard normal distribution, ϵ is the indifference interval, and γ^* is the confidence level. Bechhoffer provides extensive tables to determine d [18]. Turnbull and Weiss provide a proof that their algorithm asymptotically exhibits the requested confidence as the average variance of the hypotheses divided by the indifference interval converges to zero.

Appendix C: The COMPOSER System

The COMPOSER system [20] uses a statistical approach very similar to STOP1. Because COMPOSER performs hill climbing, it is always working from a current strategy H_0 and a candidate set of alternative strategies $H_1, \ldots, H_k$. COMPOSER computes the incremental utilities of adopting each of the alternative strategies over H_0, (i.e. COMPOSER tracks $\overline{U}_{u_1-u_0} \ldots \overline{U}_{u_k-u_0}$, computing confidence intervals for each of these distributions). COMPOSER selects n_0 samples from each distribution, then at each iteration it samples equally from each distribution. If any hypothesis $H_i \in H_1, \ldots, H_k$ is shown to have

$\overline{U}_{u_i-u_0} > 0$ with confidence γ^*, it is selected (ties are broken by the highest $\overline{U}_{u_i-u_0}$). At any iteration, any hypothesis shown to have $\overline{U}_{u_i-u_0} < 0$ with confidence γ^* is removed from the candidate set. The process terminates when a candidate strategy is selected or there are no more candidate hypotheses.

There are two major differences between COMPOSER and STOP1. First, because each strategy is compared to the default, the presence of an extremely good hypothesis strategy cannot be used to prune other hypothesis strategies. This is unfortunate because a good hypothesis strategy (e.g. better than the current strategy) can be shown to dominate a poor hypothesis more easily (faster) than the poor hypothesis can be shown to be dominated by the current strategy. The second difference is that STOP1 incorporates an indifference interval. In some cases, one or more hypotheses will have approximately the same utility as the current strategy. Thus it may take many samples to determine which strategy is better, but the overall gain or loss is insignificant. This is a poor expenditure of sampling resources.

Steve Chien is a Technical Group Leader in the Artificial Intelligence Group, Advanced Information Systems Section at the Jet Propulsion Laboratory, California Institute of Technology where he leads efforts in automated planning and scheduling. Dr. Chien is also an Adjunct Assistant Professor with the Department of Computer Science of the University of Southern California. He holds B.S., M.S., and Ph.D. degrees in Computer Science, all from the University of Illinois. His current research interests lie in the areas of: planning and scheduling, machine learning, operations research, and decision theory.

Jonathan Gratch is a Ph.D. candidate in the Computer Science Department at the University of Illinois. He holds a B.A. in Computer Science from the University of Texas at Austin and an M.S. in Computer Science from the University of Illinois. Mr. Gratch has also worked at the Army Corps of Engineers Construction Engineering Research Laboratory, where he developed artificial intelligence approaches for automated thermal system design, and at the Jet Propulsion Laboratory, where he developed learning techniques for automated scheduling. His research interests are in statistical and explanation-based learning techniques for automated planning and scheduling systems.

Michael Burl is a Ph.D. candidate in the Department of Electrical Engineering of the California Institute of Technology and a Member of Technical Staff in the Artificial Intelligence Group, Advanced Information Systems Section at the Jet Propulsion Laboratory, California Institute of Technology. He holds the B.S. degree in Applied Mathematics and Electrical Engineering and the M.S. degree in Electrical Engineering from the California Institute of Technology. From 1987 through 1991, he worked at the MIT Lincoln Laboratory in the Battlefield Surveillance Group, where he developed algorithms to detect and classify stationary military vehicles in high-resolution polarimetric Synthetic Aperature Radar (SAR) imagery. His research interests include the development of effective machine vision and learning algorithms.